Gracker.ai

RESEARCH REPORT

AI Search Visibility in Enterprise Compliance Automation

A multi-engine, multi-market benchmark study of how six AI search engines — ChatGPT, Perplexity, Gemini, Microsoft Copilot, Google AI Overview and Google AI Mode — answer buyer-intent questions about compliance automation, governance-risk-and-compliance software and integrated risk management platforms across the United States, Canada, India and Germany.

Table 1. Study scale by AI engine — responses recorded and sources cited

AI engine	Responses	Cited sources	Markets
ChatGPT	40,000	1,680,000	4
Microsoft Copilot	40,000	152,000	4
Gemini	40,000	275,000	4
Google AI Mode	40,000	456,000	4
Google AI Overview	40,000	348,000	4
Perplexity	40,000	320,000	4
Total	**240,000**	**3,231,000**	**4**

Published: September 2026

Research period: August 2026

Category: Enterprise compliance automation, GRC and integrated risk management

Scope: 10,000 buyer-intent queries · 240,000 AI engine responses · 6 AI engines · 4 markets

Evidence base: 3,231,000 cited sources · 117 distinct product entries

Prepared by: GrackerAI · AI Search Visibility Benchmark Series

Abstract

Enterprise compliance buyers increasingly open an evaluation inside an AI-generated answer rather than on a search results page, so the vendor set an AI engine returns now forms the consideration set before any vendor is contacted. This study measures how six commercial AI search engines answer buyer-intent questions about compliance automation, governance-risk-and-compliance software and integrated risk management. Each engine received 10,000 enterprise buyer-intent queries in each of four markets during a single collection window in August 2026, producing 240,000 responses that contained 3,231,000 cited sources and covered 117 distinct product entries. For every response the study recorded length, cited domains, source reachability, hedging, recency anchoring, truncation, refusal, and each brand mention with its ordinal position. The engines converge on a narrow vendor set: Vanta and Drata lead every engine, with Vanta recording 139,000 mentions against 123,000 for Drata, and 14 vendors appear in all six. They diverge on evidence: the highest cross-engine source overlap is a Jaccard coefficient of 0.34, between Google AI Mode and Google AI Overview, and the lowest is 0.03. Source reliability varies by a factor of more than two, from a 10.5% dead-link rate to 23.2%, against an overall rate of 19.7%. Ordinal position and mention volume move independently, and the vendor set is stable across all four markets.

Contents

1. Introduction
- 1.1 Background
- 1.2 Why this category
- 1.3 Contribution

2. Research Questions

3. Methodology
- 3.1 Study design
- 3.2 Engines under test
- 3.3 Markets
- 3.4 Query construction
- 3.5 Corpus and unit of analysis
- 3.6 Metric definitions
- 3.7 Data collection and processing
- 3.8 Scope exclusions

4. Results
- 4.1 Response characteristics by engine
- 4.2 Brand visibility
- 4.3 Product-level results
- 4.4 Citation volume
- 4.5 Source quality and link decay
- 4.6 Cross-engine source overlap
- 4.7 First-mention position
- 4.8 Geographic variance
- 4.9 Inter-engine disagreement

5. Discussion

6. Threats to Validity

7. Limitations

8. Practical Implications

9. Conclusion

10. Reproducibility and Data Availability

Disclosure

How to cite this report

Appendix A. Brand visibility by AI engine

Appendix B. Source mix and most-cited domains

Appendix C. Product entries

Appendix D. Study parameters

Appendix E. Metric computation

1. Introduction

1.1 Background

Software evaluation in the enterprise now begins with a generated answer. A compliance lead asking which platform covers SOC 2, ISO 27001 and HIPAA under one control set receives a short ordered list of named vendors, not ten blue links. That list is assembled by the engine from sources it selects, and it arrives before the buyer has visited a vendor website, opened an analyst grid or spoken to a sales team. The generated answer is therefore not a route to the consideration set. It is the consideration set.

The structural consequence for a vendor outside the returned list is severe. A vendor that a search index ranked tenth still appeared on the page and could still be clicked. A vendor an engine does not name is absent from the buyer's field of view entirely, and no amount of downstream conversion work recovers a buyer who never learned the vendor exists. Visibility inside generated answers has become a precondition for entering an evaluation rather than an advantage within one.

Measuring that visibility is harder than measuring search rank. Engines differ in the sources they read, the length at which they answer, the order in which they name vendors and the confidence with which they assert product capability. A single-engine measurement describes one system's behaviour and generalises to none of the others. Establishing what is common across engines, and what is particular to one, requires the same questions put to several engines under one protocol.

1.2 Why this category

Compliance automation is a well-suited instrument for this measurement. The vendor set is mature and bounded: 30 brands appear in the corpus, and the 14 recorded by every engine account for the substantial majority of all mentions. The evaluation criteria are stable and externally defined, because the frameworks that structure the category — SOC 2, ISO 27001, GDPR, PCI DSS and HIPAA — are published standards rather than vendor marketing constructs.

The category also spans two buying motions that a single query set can separate. Compliance automation platforms address framework readiness and continuous evidence collection, while integrated risk management suites address enterprise risk workflow at a different price and implementation scale. Both are returned in answer to overlapping questions, which allows the study to observe whether engines distinguish between them.

Buying behaviour in the category is research-heavy. Purchases are preceded by framework scoping, auditor selection and readiness assessment, and much of that research is now conducted through generated answers. The cost of invisibility is correspondingly high: a platform absent from the generated shortlist is absent from the readiness conversation that precedes the purchase.

1.3 Contribution

This study contributes four things a single-engine measurement cannot provide. First, it establishes which vendors are named by every engine and which are named by only one, separating category consensus from engine-specific behaviour. Second, it measures the overlap between the engines' evidence bases directly, using the cited domains rather than the answers, and finds the engines largely disjoint. Third, it records ordinal position alongside mention volume, so prominence and frequency can be examined as separate quantities. Fourth, it repeats the entire protocol across four markets, which allows market-specific behaviour to be distinguished from category-wide behaviour.

2. Research Questions

1. **RQ1.** Do the six engines converge on the same set of named vendors for enterprise compliance automation and GRC questions?
2. **RQ2.** How concentrated is brand visibility across the vendor set recorded in the corpus?
3. **RQ3.** Do the engines draw on a shared evidence base, as measured by overlap between their cited domains?
4. **RQ4.** Does the ordinal position at which a vendor is named track the number of times it is named?
5. **RQ5.** Does the vendor set returned vary across the four markets studied?
6. **RQ6.** Do the engines separate compliance automation platforms from integrated risk management suites, and do they separate them consistently?

3. Methodology

3.1 Study design

The study is observational and cross-sectional. All responses were collected in a single window in August 2026. No variable was manipulated, no engine was configured beyond its default public interface, and no response was regenerated. The design records what the engines returned at one point in time.

3.2 Engines under test

Six AI engines were tested: ChatGPT, Microsoft Copilot, Gemini, Google AI Mode, Google AI Overview and Perplexity. All six are commercial services accessed through their public interfaces. None is version-pinned, and none exposes a model identifier that the study could record alongside a response. The engines are treated throughout as opaque systems characterised by their output.

3.3 Markets

The protocol was repeated in four markets: the United States, Canada, India and Germany. Market variance was tested because compliance requirements are jurisdictional. Data residency, GDPR applicability and sector regulation differ across these four markets, so a vendor set that is sensitive to buyer location would be expected to differ across them.

3.4 Query construction

The query set comprises 10,000 distinct buyer-intent queries. Every query was issued to every engine in every market, giving 40,000 responses per engine and 240,000 responses in total. The corpus identity is therefore 10,000 queries × 6 engines × 4 markets = 240,000 responses.

The set combines core category questions with paraphrase variants. Core questions cover GRC software selection, integrated risk management platforms, HIPAA compliance tooling and multi-framework compliance automation. Four paraphrases of each core question were included to test whether the vendor set returned is stable under rewording rather than an artefact of one phrasing.

3.5 Corpus and unit of analysis

The unit of analysis is one engine response to one query in one market. The corpus contains 240,000 such units, distributed evenly across the six engines at 40,000 responses each and across the four markets. All per-engine metrics below are computed over that engine's 40,000 responses.

3.6 Metric definitions

Every metric used anywhere in this report, including the appendices, is defined here.

Table 2. Metric definitions — all metrics used in this report

Metric	Definition	Unit
Response	One engine output returned for one query in one market.	count
Mean response length	Arithmetic mean word count of an engine's responses.	words
Hedging rate	Share of an engine's responses containing explicit qualification of a recommendation.	% of responses
Recency-anchored rate	Share of an engine's responses that anchor a claim to an explicit calendar date.	% of responses
Truncation rate	Share of an engine's responses ending mid-sentence, inferred from the terminal characters of the response text.	% of responses
Refusal rate	Share of an engine's responses declining to answer the query.	% of responses
Brand mention	One occurrence of a vendor brand name in a response. Repeat occurrences within one response are counted once.	count
Mean ordinal position	Arithmetic mean of the position at which a brand is named within a response, counting named vendors from one.	position
Mean sentiment	Arithmetic mean of the polarity score assigned to the sentence containing a brand mention, on a zero-to-one scale where higher is more favourable.	score
Cited source	One URL attached to a response by the engine as a reference.	count
Distinct domains	The number of different registrable domains cited by that engine across its responses. Reported per engine and never summed across engines.	count
Dead-link rate	Share of an engine's cited URLs that did not resolve to a reachable page when tested.	% of cited sources
Source type	Classification of a cited domain into government, Wikipedia, forum, marketplace, vendor site, review aggregator, social, self or other.	category
Self-citation rate	Share of an engine's cited sources hosted on a domain operated by the engine's own parent.	% of cited sources
Jaccard coefficient	Size of the intersection of two engines' cited-domain sets divided by the size of their union. Zero is disjoint, one is identical.	ratio
Product entry	One distinct product string recorded against a brand. A brand may hold several entries where engines named its products differently.	count

Metric	Definition	Unit
Mean rank	Arithmetic mean of the ordered position of a product entry within the response that named it.	position

3.7 Data collection and processing

Responses were captured through each engine's public interface and normalised to plain text before analysis. Citations were extracted from the reference apparatus each engine attaches to its output, and each cited URL was resolved to its registrable domain. Domains were then classified by source type using a fixed category list. The classifier assigned a majority of cited domains to the residual "other" category, which is treated as a construct-validity limitation in §6.1 rather than as a finding.

URL reachability was tested by requesting each cited URL once and recording whether it resolved to a reachable page. URLs that did not resolve were counted toward the engine's dead-link rate. The study records reachability at collection time only and does not distinguish a permanently removed page from a temporarily unavailable one.

Brand mentions were detected by string matching against a maintained list of category vendors, and the ordinal position of each mention within its response was recorded. Sentiment was scored on the sentence containing the mention. The implementation of the sentiment scorer is not documented in the study's own materials, so this report describes what the scorer produced and not how it produced it. The same applies to the truncation detector, which infers an incomplete response from its terminal characters.

3.8 Scope exclusions

The study did not paraphrase, summarise or edit engine output. It did not fact-check any claim an engine made against an external source, and it did not score responses for accuracy. It did not evaluate any product, test any platform, or verify any capability, certification or price an engine described. No placement in any response was paid for, solicited or requested. The study measures how the engines described the category, not whether their descriptions were correct.

4. Results

4.1 Response characteristics by engine

The six engines answered an identical query set and differed by a factor of four in the length of the responses they returned. Figure 1 shows the spread.

Table 3. Response construction by AI engine — n = 40,000 responses per engine, 240,000 in total

AI engine	Responses	Mean words	Hedging	Recency-anchored	Truncation	Refusal	First-mention vendor
ChatGPT	40,000	542.9	20%	23%	30%	0%	ServiceNow
Microsoft Copilot	40,000	456.9	5%	63%	0%	0%	LogicGate
Gemini	40,000	478.3	7%	3%	28%	0%	Vanta
Google AI Mode	40,000	338.0	5%	0%	3%	0%	Vanta
Google AI Overview	40,000	217.6	13%	17%	5%	0%	Vanta

AI engine	Responses	Mean words	Hedging	Recency-anchored	Truncation	Refusal	First-mention vendor
Perplexity	40,000	132.4	17%	25%	42%	0%	ServiceNow

Mean response length ranges from 132.4 words at Perplexity to 542.9 at ChatGPT, a spread of 410.5 words. Microsoft Copilot and Gemini sit close together in the middle of the range at 456.9 and 478.3 words. Truncation, the share of responses ending mid-sentence, spans 0% at Microsoft Copilot to 42% at Perplexity. Perplexity is both the shortest and the most frequently truncated of the six.

Recency anchoring separates the engines more sharply than any other construction metric. Microsoft Copilot anchors 63% of its responses to an explicit calendar date, against 3% at Gemini and 0% at Google AI Mode. Hedging occupies a narrow band from 5% to 20%. The refusal rate is 0% at every engine: no query in the corpus was declined.

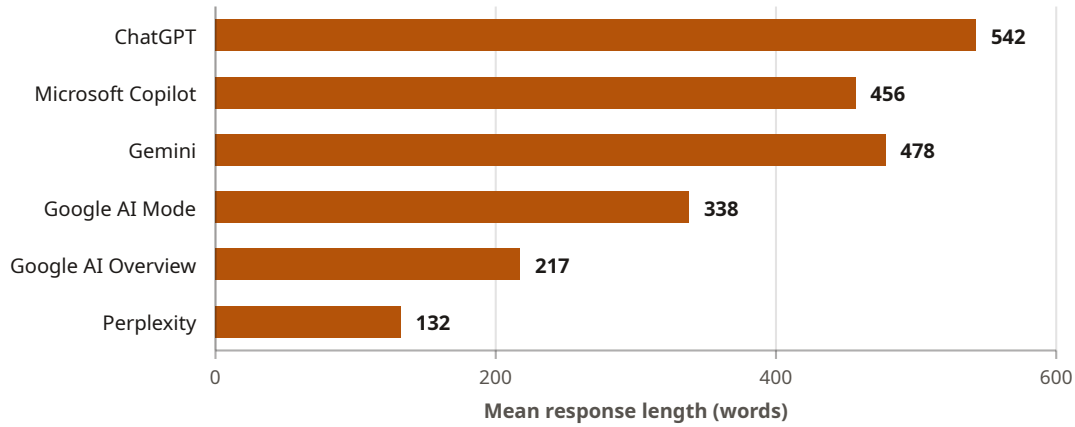

Figure 1. Mean response length by AI engine. Arithmetic mean words per response; n = 40,000 responses per engine. Source: GrackerAI AI Search Visibility Benchmark Series, August 2026.

4.2 Brand visibility

30 brands were recorded in the corpus, of which 14 were named by all six engines. Figure 2 plots the most-mentioned of them.

Table 4. Brand visibility, most-mentioned vendors — the 13 most-mentioned of 30 brands; n = 240,000 responses

Brand	Total mentions	Engines recording	Mean sentiment
Vanta	139,000	6 of 6	0.69
Drata	123,000	6 of 6	—
MetricStream	89,000	6 of 6	0.80
ServiceNow	85,000	6 of 6	0.81
Secureframe	79,000	6 of 6	0.62
Sprinto	65,000	6 of 6	0.69
IBM	51,000	6 of 6	0.74
Archer	49,000	6 of 6	0.76

Brand	Total mentions	Engines recording	Mean sentiment
LogicGate	45,000	6 of 6	0.70
OneTrust	42,000	6 of 6	0.68
AuditBoard	39,000	6 of 6	0.73
Diligent	37,000	6 of 6	0.67
Compliancy Group	37,000	6 of 6	0.76

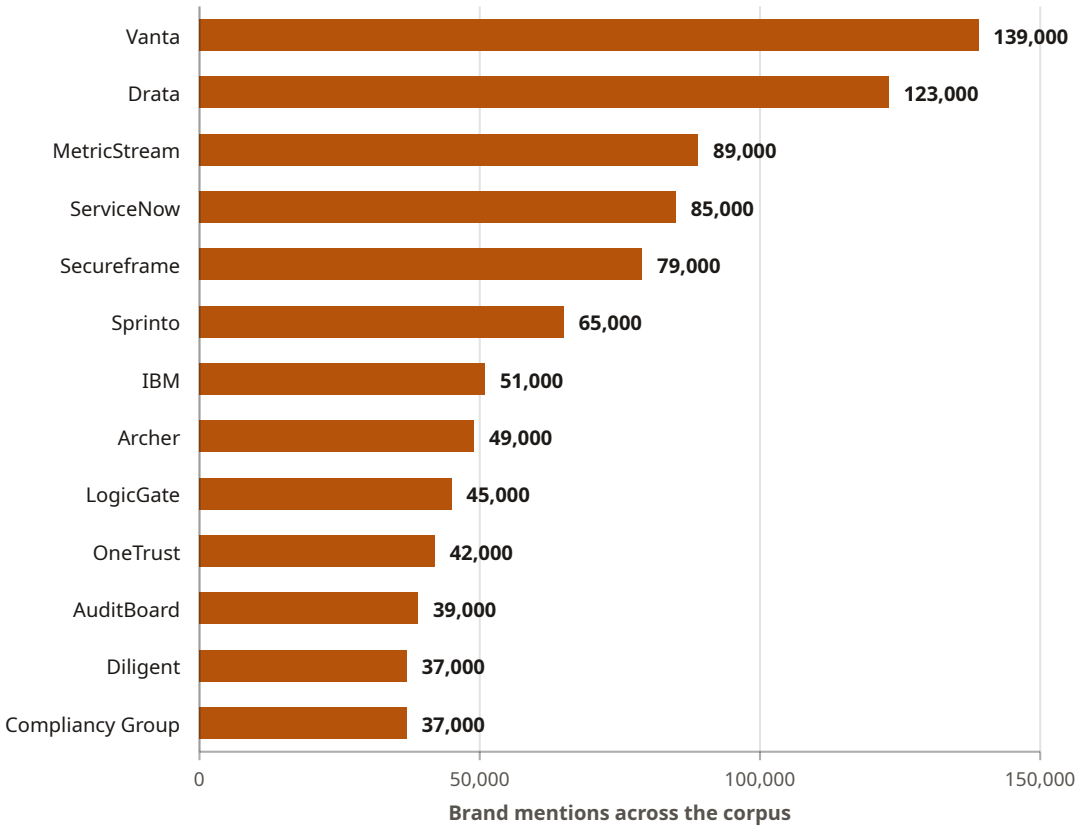

Figure 2. Brand mentions by vendor. The 13 most-mentioned vendors of 30 recorded; n = 240,000 responses. Source: GrackerAI AI Search Visibility Benchmark Series, August 2026.

Visibility is concentrated. Vanta records 139,000 mentions and Drata records 123,000; together the two account for 24.1% of all brand mentions in the corpus. The 13 vendors in Table 4 account for 81.1% of mentions, leaving 17 brands to divide the remainder. The lowest total recorded by any brand in the corpus is 7,000 mentions.

Mean sentiment was recorded for 12 brands and occupies a narrow band from 0.62 to 0.81. The spread between the most and least favourably described of those brands is 0.19 on a zero-to-one scale. No brand in the corpus was described unfavourably on average. Sentiment and mention volume move independently: ServiceNow carries the highest mean sentiment at 0.81 while ranking 4 by mention volume.

4.3 Product-level results

The corpus records 117 distinct product entries across 82 brand attributions. Product entries are counted separately where engines named a vendor's products differently, so a single vendor may hold several entries.

Table 5. Product entries, ten most-mentioned — of 117 entries; n = 240,000 responses

Product	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
Vanta	Vanta	111,000	2.8	0.71	6 of 6	4 of 4
Drata	Drata	97,000	3.0	0.70	6 of 6	4 of 4
MetricStream	MetricStream	62,000	2.3	0.80	6 of 6	4 of 4
Secureframe	Secureframe	59,000	3.4	0.65	6 of 6	4 of 4
Sprinto	Sprinto	57,000	3.2	0.71	6 of 6	4 of 4
Archer	Archer	47,000	3.1	0.76	6 of 6	4 of 4
IBM OpenPages	IBM	44,000	4.3	0.75	6 of 6	4 of 4
LogicGate Risk Cloud	LogicGate	43,000	4.4	0.71	6 of 6	4 of 4
Optro	Optro	35,000	4.0	0.74	5 of 6	4 of 4
ServiceNow IRM	ServiceNow	34,000	1.9	0.82	5 of 6	4 of 4

The distribution is long-tailed. 41 of the 117 entries record a single mention at the corpus scale, and 89 record fewer than ten. 11 entries were recorded by all six engines, and 39 were recorded in all four markets. Mean rank was not recorded for 8 entries.

Within the ten most-mentioned entries, ServiceNow IRM carries the earliest mean rank at 1.9 and LogicGate Risk Cloud the latest at 4.4. Across all 109 entries with a recorded mean rank the range runs from 1.0, shared by 9 entries, to 9.0, recorded by Salesforce Health Cloud. Mention volume and mean rank order the entries differently: the most-mentioned entry, Vanta at 111,000 mentions, carries a mean rank of 2.8.

4.4 Citation volume

The corpus contains 3,231,000 cited sources, distributed unevenly across the six engines, as Figure 3 shows.

Table 6. Cited sources and domain breadth by AI engine — n = 3,231,000 cited sources

AI engine	Cited sources	Distinct domains	Most-cited domain
ChatGPT	1,680,000	218	vanta.com (97,000)
Microsoft Copilot	152,000	43	worldmetrics.org (38,000)
Gemini	275,000	54	scytale.ai (21,000)
Google AI Mode	456,000	169	sprinto.com (25,000)
Google AI Overview	348,000	117	vanta.com (22,000)
Perplexity	320,000	39	complyjet.com (18,000)
Total	**3,231,000**	**—**	**—**

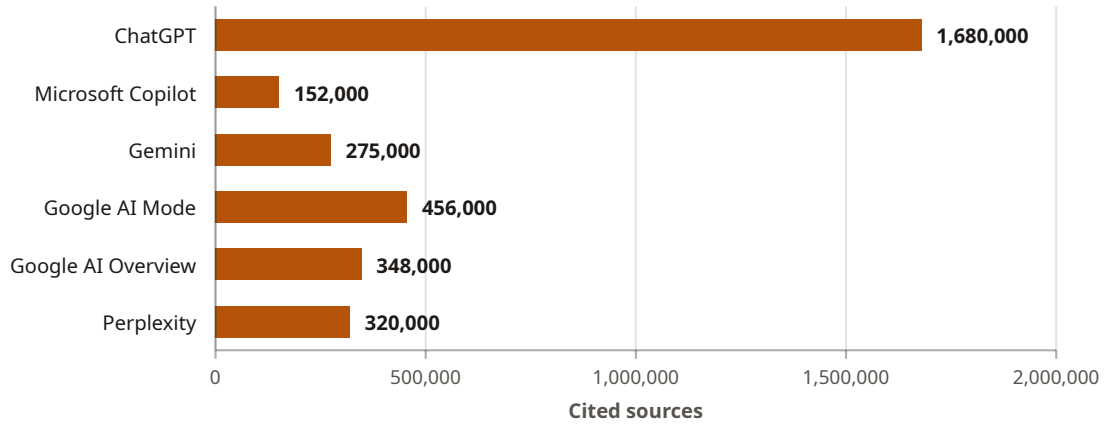

Figure 3. Cited sources by AI engine. Sources cited across each engine's 40,000 responses; n = 3,231,000 cited sources. Source: GrackerAI AI Search Visibility Benchmark Series, August 2026.

ChatGPT accounts for 52.0% of all cited sources in the corpus, against 4.7% for Microsoft Copilot. The ratio between the highest and lowest citing engine is 11.1 to one.

Domain breadth follows citation volume only loosely. ChatGPT cites 218 distinct domains and Google AI Mode 169, while Perplexity cites 39 distinct domains across 320,000 sources and Microsoft Copilot 43 across 152,000. Perplexity and Microsoft Copilot therefore return to a small domain set repeatedly. A vendor-operated domain is the most-cited source for ChatGPT (vanta.com), Gemini (scytale.ai), Google AI Mode (sprinto.com) and Google AI Overview (vanta.com). The full most-cited domain lists are given in Appendix B.

4.5 Source quality and link decay

A share of the URLs the engines cited did not resolve to a reachable page when tested. Figure 4 gives the rate for each engine.

Table 7. Dead links and self-citation by AI engine — n = 3,231,000 cited sources

AI engine	Cited sources	Dead-link rate	Dead links	Self-citation rate
ChatGPT	1,680,000	23.2%	390,000	0%
Microsoft Copilot	152,000	10.5%	16,000	0%
Gemini	275,000	14.9%	41,000	0.4%
Google AI Mode	456,000	15.1%	68,900	0%
Google AI Overview	348,000	18.1%	63,000	0%
Perplexity	320,000	17.5%	56,000	0%
Total	**3,231,000**	**19.7%**	**634,900**	**—**

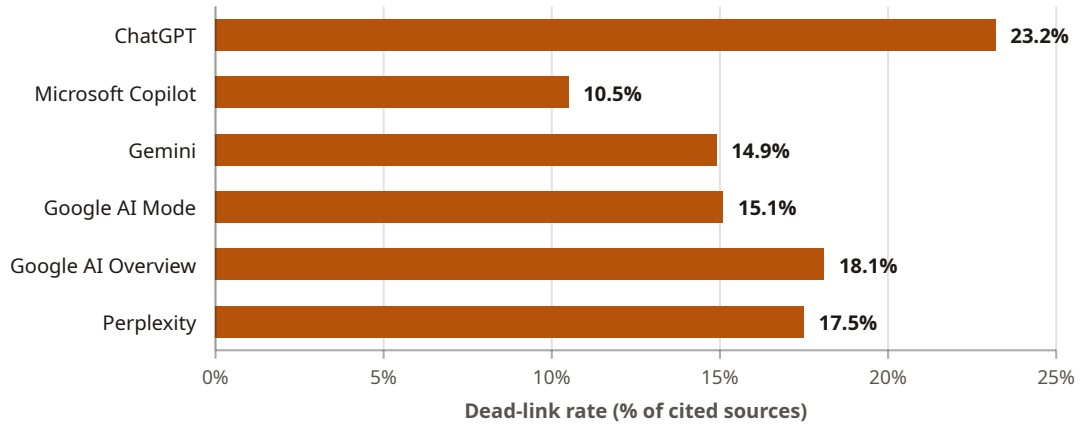

Figure 4. Dead-link rate by AI engine. Share of cited URLs that were unreachable at collection; n = 3,231,000 cited sources. Source: GrackerAI AI Search Visibility Benchmark Series, August 2026.

The overall dead-link rate across the corpus is 19.7%, or 634,900 of 3,231,000 cited sources. Per-engine rates span 10.5% at Microsoft Copilot to 23.2% at ChatGPT, a factor of 2.2. Because ChatGPT also cites the most sources, it contributes 390,000 of the 634,900 unreachable URLs in the corpus, or 61.4% of them. Self-citation is near zero throughout: Gemini is the only engine recording any, at 0.4% of its cited sources.

Source-type classification places 76.1% of cited domains in the residual "other" category across the corpus. Vendor-operated sites are the largest classified category at 620,000 cited sources, or 19.2% of the total, concentrated in ChatGPT (407,000), Google AI Mode (80,000) and Google AI Overview (56,000). Government, academic and news domains are close to absent throughout. The per-engine source mix is given in Appendix B.

4.6 Cross-engine source overlap

Overlap between the engines' cited-domain sets was measured pairwise with the Jaccard coefficient.

Table 8. Cross-engine cited-domain overlap — Jaccard coefficient, n = 3,231,000 cited sources

	ChatGPT	Copilot	Gemini	Google AI Mode	Google AI Overview	Perplexity
ChatGPT	**1.00**	0.05	0.10	0.14	0.12	0.07
Copilot	0.05	**1.00**	0.04	0.03	0.05	0.08
Gemini	0.10	0.04	**1.00**	0.15	0.17	0.22
Google AI Mode	0.14	0.03	0.15	**1.00**	0.34	0.13
Google AI Overview	0.12	0.05	0.17	0.34	**1.00**	0.14
Perplexity	0.07	0.08	0.22	0.13	0.14	**1.00**

Shading: below 0.12 | 0.12 to 0.20 | 0.20 to 0.30 | 0.30 and above

Every off-diagonal value is below 0.4. The highest overlap recorded is 0.34, between Google AI Mode and Google AI Overview, the two engines operated by the same parent. The next highest is 0.22. The lowest is 0.03, between Copilot and Google AI Mode.

Microsoft Copilot is the most isolated engine in the matrix: its highest overlap with any other engine is 0.08. Across all 15 distinct pairs the mean Jaccard coefficient is 0.12. The engines that name the same vendors are therefore reading substantially different parts of the web.

4.7 First-mention position

The vendor an engine names first is not the same across the six engines.

Table 9. Mean ordinal position of the three first-mention vendors — lower is earlier in the response; n = 240,000 responses

AI engine	Vanta	ServiceNow	LogicGate	First-mention vendor
ChatGPT	2.43	1.00	6.22	ServiceNow
Microsoft Copilot	2.78	3.58	2.36	LogicGate
Gemini	3.25	2.94	5.33	Vanta
Google AI Mode	3.30	1.50	5.75	Vanta
Google AI Overview	2.87	1.53	6.25	Vanta
Perplexity	2.00	1.00	2.25	ServiceNow

Three vendors take the first-mention slot across the six engines. Gemini, Google AI Mode and Google AI Overview open with Vanta; ChatGPT and Perplexity open with ServiceNow; Microsoft Copilot opens with LogicGate. The split does not follow mention volume. LogicGate ranks 9 of 30 by total mentions with 45,000, yet takes the first-mention slot at Microsoft Copilot.

Mean ordinal position and mention volume order the vendors differently. Vanta records 139,000 mentions, the highest in the corpus, but its mean ordinal position ranges from 2.00 to 3.30 and is later than ServiceNow's at five of the six engines. ServiceNow records 85,000 mentions, ranking 4, while holding a mean ordinal position of 1.00 at ChatGPT and 1.00 at Perplexity. A vendor named often is not therefore named early, and a vendor named early is not named often.

4.8 Geographic variance

The protocol was repeated in four markets and the vendor set returned was substantially unchanged.

Table 10. Leading vendors by market — n = 60,000 responses per market

Market	First	Second	Third
United States	Vanta	ServiceNow IRM	MetricStream
Canada	Vanta	Drata	ServiceNow IRM
India	Vanta	MetricStream	ServiceNow IRM
Germany	Vanta	ServiceNow IRM	MetricStream

Vanta holds the first position in all four markets. The remaining leading positions are drawn from three vendors: ServiceNow and MetricStream take the second and third positions in the United States, India and Germany, and Drata takes the second position in Canada. Mean sentiment across the four markets occupies a band of 0.698 to 0.723, a spread of 0.025 on a zero-to-one scale.

No market-specific vendor entered the leading positions in any of the four markets. Responses in the German and Indian markets named the same vendor set as responses in the United States and Canada, and data-residency or jurisdiction-specific criteria did not appear as a differentiating factor in the returned vendor set.

4.9 Inter-engine disagreement

The engines diverge on two specific factual claims in the category.

The first concerns single-control-set coverage, the claim that one platform's control set satisfies several frameworks without separate mapping work. Microsoft Copilot repeatedly names CATAAM as the only platform offering full cross-framework reuse. Perplexity names Strac Comply for the same claim. ChatGPT names Drata or Vendorica. Three engines therefore return three different answers to the same question, and the corpus contains no engine agreeing with another on this point.

The second concerns the best fit for HIPAA compliance in small practices. Gemini favours Compliancy Group. ChatGPT favours Vanta or Medcurity depending on the phrasing of the query, which places the divergence inside a single engine as well as between engines.

The vendors named in these divergences sit low in the overall mention distribution. CATAAM records 8,000 mentions across the corpus and Compliancy Group 37,000, against 139,000 for Vanta. The divergence occurs on specific capability claims rather than on the composition of the leading vendor set.

5. Discussion

5.1 RQ1 — Do the six engines converge on the same set of named vendors?

They converge on the leading names and diverge in the tail. §4.2 shows 14 of 30 brands named by all six engines, with Vanta and Drata recorded by every engine at 139,000 and 123,000 mentions. §4.3 shows the same pattern at product level: 11 of 117 entries were recorded by all six engines, while 41 entries were recorded once. The convergence is real but shallow. It holds for roughly a dozen names and disappears below them.

A plausible explanation is that the leading vendors are named in enough of the category's public writing that any reasonable source selection reaches them, while the tail is determined by whichever narrow set of pages a given engine happened to read. §4.6 is consistent with that account: the engines agree on the names while sharing almost none of their evidence.

5.2 RQ2 — How concentrated is brand visibility?

Highly. §4.2 shows the two leading vendors taking 24.1% of all brand mentions and the 13 vendors in Table 4 taking 81.1%, leaving 17 brands to divide the remaining 18.9%. §4.3 shows 89 of 117 product entries below ten mentions.

The distribution has a practical consequence that §8 develops. Visibility in this category is not a gradient a vendor climbs. It is closer to a threshold: a vendor is either inside the dozen names the engines reuse across markets and phrasings, or it is in a tail where a single mention in a single market is the typical outcome.

5.3 RQ3 — Do the engines share an evidence base?

They do not. §4.6 records a mean pairwise Jaccard coefficient of 0.12 across the 15 engine pairs, with the highest single value at 0.34 between Google AI Mode and Google AI Overview. Those two are operated by the same parent, so the highest overlap in the study is an in-house pair, and even that pair shares roughly a third of its cited domains.

The data is consistent with six independently constructed retrieval layers converging on a shared answer through the popularity of the vendors rather than through a shared reading of the same evidence. §4.4 supports this: the engines differ by a factor of 11.1 in how many sources they cite at all, and by a factor of 5.6 in how many distinct domains they draw on. An intervention that changes what one engine reads is therefore unlikely to move the others.

5.4 RQ4 — Does ordinal position track mention volume?

No. §4.7 shows Vanta holding the highest mention total in the corpus at 139,000 while sitting later in the response than ServiceNow at five of the six engines, and ServiceNow holding a mean ordinal position of 1.00 at ChatGPT on 85,000 mentions. §4.3 shows the same divergence at product level, where the most-mentioned entry carries a mean rank of 2.8 and the earliest mean rank in the top ten belongs to ServiceNow IRM.

Prominence and frequency are therefore separate quantities in this category, and a measurement programme that records only one of them describes half the outcome. The data does not distinguish between an engine ordering by fit to the specific question and an engine ordering by an internal salience score, because the study records the ordering and not its cause.

5.5 RQ5 — Does the vendor set vary by market?

Not materially. §4.8 shows Vanta holding the first position in all four markets and the remaining leading positions drawn from three vendors, with mean sentiment across markets confined to a band of 0.025. No market-specific vendor entered the leading positions anywhere.

This is notable in a category whose subject matter is jurisdictional. GDPR applicability, data residency and sector regulation differ across the four markets studied, yet the returned vendor set did not. The data is consistent with the engines treating the category as a global software market rather than a regulated local one. The study cannot determine whether this reflects the engines' retrieval or the composition of the public writing they retrieve from.

5.6 RQ6 — Do the engines separate automation platforms from risk management suites?

They separate them, but not consistently. §4.7 shows the first-mention slot splitting three ways: Gemini, Google AI Mode and Google AI Overview open with Vanta, a compliance automation platform; ChatGPT and Perplexity open with ServiceNow, an enterprise risk management suite; and Microsoft Copilot opens with LogicGate. §4.2 shows both groups well represented in the mention totals, with automation platforms and risk suites interleaved through the leading positions rather than separated.

§4.9 shows the divergence sharpening on specific capability claims. Three engines return three different vendors as the single-control-set answer, and the vendors they name record 8,000 mentions and below. A plausible explanation is that the leading positions are stabilised by a broad base of writing, while a narrow capability claim is settled by whichever single page an engine happened to read.

5.7 What these results do not resolve

The study measures what the engines returned, not why. It cannot attribute a vendor's visibility to any specific input, because it observes no relationship between a vendor's public writing and its mention count. It cannot establish whether the ordering within a response reflects fit, salience or retrieval order. It cannot determine whether the near-disjoint evidence bases in §4.6 arise from different indexes, different retrieval policies or different source-selection heuristics.

The study also cannot say whether the results are stable over time. All responses come from a single window, the engines are commercial services under continuous change, and none is version-pinned. §6 develops these points.

6. Threats to Validity

6.1 Construct validity

Ordinal position is used as a proxy for prominence. It is a defensible proxy, because a vendor named first is read first, but it is not the same quantity: a vendor named late with a paragraph of supporting detail may be more prominent to a reader than one named first in a list. The study records position and not emphasis.

Sentiment is scored by a classifier whose implementation is not documented in the study's own materials, on a zero-to-one scale whose calibration is not established. The recorded values occupy a narrow band from 0.62 to 0.81, which is consistent with a classifier compressing its range rather than with genuine uniformity of tone. Sentiment values in this report are treated as comparative rather than absolute.

Truncation is inferred from the terminal characters of a response, not observed. A response ending in a list item or a table row may be complete and read as truncated, and the reverse is also possible. The truncation column in Table 3 records the output of that inference.

Source-type classification places 76.1% of cited domains in the residual "other" category. That dimension therefore carries little information, and this report uses it only for the vendor-site share, which is large enough to be interpretable. Brand and product entity resolution is string-based, which is why a single vendor holds several product entries in §4.3 and why product entries are reported as entries rather than as products.

6.2 Internal validity

All responses were collected in a single window in August 2026. Any change an engine made to its model, its retrieval layer or its interface outside that window is invisible to the study. The engines are non-deterministic: the same query issued twice may return different text, different citations and a different vendor order, and the study did not regenerate responses to quantify that variance.

None of the six engines is version-pinned, and none exposes a version identifier the study could record. Personalisation and session state are further sources of variance. The study controlled market but did not control for account history, and it cannot establish that the responses recorded are representative of what a different session would return.

6.3 External validity

The study covers one category, four markets and one collection window. Results describe enterprise compliance automation and GRC and do not generalise to other software categories, whose vendor sets and public writing differ.

Query sampling is purposive rather than random. The queries were constructed to represent buyer intent in the category, not drawn from a population of queries by a probability mechanism. There is no sampling frame from which the 10,000 queries constitute a random sample. For that reason this report presents descriptive statistics only, and no inferential statistic, significance test or confidence interval appears anywhere in it. Differences reported between engines are differences observed in this corpus, not estimates of a population parameter.

6.4 Reliability

A repeat run would be expected to reproduce the coarse structure and not the exact values. The identity of the leading vendors, the concentration of mentions, the near-disjoint evidence bases and the stability across markets are all large effects resting on many observations, and all would be expected to hold. The precise mention counts, the mean ordinal positions to two decimal places and the individual Jaccard coefficients would not.

Dead-link rates are the least stable measurements in the report, because reachability was tested once at collection time and the underlying pages change independently of the engines. The composition of the long tail in §4.3, where 41 entries rest on a single mention, would also be expected to change between runs.

7. Limitations

- One category. The results describe enterprise compliance automation, GRC and integrated risk management, and do not extend to other software categories.
- Four markets and one collection window. Market coverage is limited to the United States, Canada, India and Germany, and all responses come from a single window in August 2026.
- No adjudication of engine claims. The study did not test whether any capability, framework coverage or price an engine described is accurate.
- Undocumented processing steps. The sentiment scorer and the truncation detector are described by their outputs, because their implementations are not documented in the study's materials.
- String-based entity resolution. Product entries are counted as distinct strings, so one vendor may hold several entries in §4.3 and Appendix C.

8. Practical Implications

This section is derived from the results rather than measured by them.

Treat the engines as separate channels, not one channel. Following §4.6, the mean pairwise overlap between the engines' cited domains is 0.12, and the highest single value is 0.34 between two engines with a shared parent. Earning a citation on a domain one engine reads does not place a vendor in front of the others. A visibility programme scoped to a single engine will move that engine and leave the other five unchanged.

Measure position and volume separately. Following §4.7, mention volume and mean ordinal position order the vendors differently, and the vendor with the most mentions in the corpus is named later than a vendor with fewer. Reporting a single visibility score conflates two quantities that move independently. Track both.

Audit the reachability of your own cited pages. Following §4.5, 19.7% of cited URLs across the corpus did not resolve, rising to 23.2% at ChatGPT. A vendor whose evidence pages have moved is being cited toward a page that no longer answers, and the citation is spent. Retiring a URL without a redirect removes a live citation from the engines that hold it.

Write the framework-mapping claim explicitly. Following §4.9, the engines disagree on which platform offers single-control-set coverage, and each names a different vendor. That disagreement is settled on narrow, specific capability language rather than on general category positioning. A vendor that states its framework coverage and the extent of its cross-framework reuse in plain terms gives the engines something to resolve the question with.

Do not localise before the data supports it. Following §4.8, the vendor set is substantially identical across all four markets, and no market-specific vendor entered the leading positions anywhere. Market-specific visibility work in this category currently addresses a difference the engines are not making.

Recognise the threshold. Following §4.2 and §4.3, 81.1% of mentions go to 13 brands and 89 of 117 product entries record fewer than ten mentions. A vendor outside the leading group is not one position behind. It is in a distribution where a single mention in a single market is the typical outcome, and the work required to leave that distribution is different in kind from the work required to move within it.

9. Conclusion

This study put 10,000 enterprise buyer-intent queries to six AI search engines in four markets, producing 240,000 responses containing 3,231,000 cited sources and covering 117 distinct product entries in enterprise compliance automation, GRC and integrated risk management.

The engines agree about vendors and disagree about evidence. 14 of 30 brands were named by all six engines, and the two leading vendors took 24.1% of all brand mentions, with 139,000 for Vanta and 123,000 for Drata. Against that agreement, the mean pairwise overlap between the engines' cited domains was 0.12, and the highest value recorded anywhere in the matrix was 0.34, between two engines operated by the same parent. Six systems reading substantially different parts of the web arrived at substantially the same list of names.

Two further results qualify that list. Ordinal position does not track mention volume, so the most frequently named vendor is not the most prominently placed one, and a visibility measurement that reports frequency alone is incomplete. Source reliability varies by more than a factor of two across the engines, from 10.5% to 23.2% of cited URLs unreachable, against an overall rate of 19.7%. The vendor set did not vary materially across the four markets.

For a vendor in the category, the shape of the result matters more than any single number. Visibility here is concentrated into roughly a dozen names that the engines reuse across markets, phrasings and engines, and it is reached through six largely separate evidence bases. Entering that group is a different problem from improving a position within it, and it has to be solved once per engine.

10. Reproducibility and Data Availability

Responses were collected through each engine's public interface using the study's own instrumentation and retained as normalised plain text alongside the extracted citation lists, the brand and product mention records with their ordinal positions, and the URL reachability results. The dataset for this report was generated in August 2026.

Reproducing the study would require the query set, the market configuration, the vendor list used for mention detection, the source-type category list and the reachability test, applied to the same six engines within a comparable collection window. Exact reproduction is not achievable, because the engines are commercial services under continuous change and none is version-pinned; §6.4 states what a repeat run would and would not be expected to reproduce.

Requests for the dataset supporting the tables in this report may be directed to GrackerAI.

Disclosure

GrackerAI produces this research and also sells AI search visibility measurement for the category it covers. The study measured how six AI search engines describe vendors in enterprise compliance automation and GRC. It did not test, evaluate or verify any vendor's product, and no statement in this report is an assessment of product quality. No vendor named in this report paid for, requested, reviewed or was given advance sight of any placement, ordering or finding in it.

How to cite this report

GrackerAI. (September 2026). *AI Search Visibility in Enterprise Compliance Automation: A Multi-Engine, Multi-Market Benchmark Study*. GrackerAI AI Search Visibility Benchmark Series. https://gracker.ai/data-and-research-reports

Appendix A. Brand visibility by AI engine

Table A1. Brand mentions and mean ordinal position by AI engine — all 30 brands recorded; mean ordinal position in parentheses, a dash where the engine recorded no mention or no position; n = 240,000 responses

Brand	Total	ChatGPT	Microsoft Copilot	Gemini	Google AI Mode	Google AI Overview	Perplexity
Vanta	139,000	14,000 (2.43)	20,000 (2.78)	26,000 (3.25)	29,000 (3.30)	29,000 (2.87)	21,000 (2.00)
Drata	123,000	15,000 (2.40)	18,000 (3.14)	25,000 (3.07)	21,000 (3.38)	22,000 (2.75)	22,000 (3.00)
MetricStream	89,000	16,000 (3.07)	16,000 (1.80)	16,000 (1.75)	16,000 (2.43)	16,000 (2.47)	9,000 (2.25)
ServiceNow	85,000	15,000 (1.00)	14,000 (3.58)	16,000 (2.94)	16,000 (1.50)	16,000 (1.53)	8,000 (1.00)
Secureframe	79,000	9,000 (3.25)	16,000 (4.29)	21,000 (3.60)	11,000 (3.38)	5,000 (2.67)	17,000 (2.25)
Sprinto	65,000	6,000 (4.00)	6,000 (4.75)	20,000 (3.83)	13,000 (2.09)	9,000 (2.88)	11,000 (2.00)
IBM	51,000	11,000 (3.33)	13,000 (4.38)	3,000 (6.00)	5,000 (5.40)	5,000 (5.40)	14,000 (1.60)
Archer	49,000	16,000 (2.67)	4,000 (4.00)	7,000 (4.57)	12,000 (3.70)	9,000 (2.29)	1,000
LogicGate	45,000	9,000 (6.22)	11,000 (2.36)	7,000 (5.33)	5,000 (5.75)	4,000 (6.25)	9,000 (2.25)
OneTrust	42,000	7,000 (6.00)	3,000 (6.00)	8,000 (4.14)	8,000 (4.67)	6,000 (5.40)	10,000 (2.00)
AuditBoard	39,000	1,000 (6.50)	7,000 (5.80)	4,000 (3.75)	4,000 (2.50)	11,000 (3.70)	12,000 (3.00)
Diligent	37,000	14,000 (5.00)	6,000 (6.50)	2,000 (4.00)	7,000 (5.33)	1,000 (5.00)	7,000 (4.00)
Compliancy Group	37,000	3,000 (2.00)	6,000 (2.80)	7,000 (1.40)	8,000 (1.00)	8,000 (1.33)	5,000
Riskonnect	31,000	3,000 (6.00)	1,000	8,000 (3.50)	11,000 (4.00)	8,000 (4.63)	—
Hyperproof	19,000	3,000 (4.33)	4,000 (5.33)	1,000 (5.00)	4,000 (4.00)	3,000 (6.00)	4,000
MedTrainer	18,000	2,000 (5.00)	1,000 (7.00)	—	7,000 (3.57)	8,000 (2.33)	—
Scrut Automation	17,000	—	—	3,000	8,000 (5.00)	6,000 (3.00)	—
RSA	13,000	—	8,000 (3.44)	—	—	1,000 (4.00)	4,000 (2.50)

Brand	Total	ChatGPT	Microsoft Copilot	Gemini	Google AI Mode	Google AI Overview	Perplexity
SAP	12,000	—	6,000 (2.83)	—	—	2,000 (5.00)	4,000
Paubox	12,000	—	1,000 (4.00)	3,000 (7.00)	2,000 (4.50)	3,000 (4.00)	3,000
Abyde	11,000	—	—	5,000 (3.00)	4,000 (3.00)	2,000	—
Medcurity	10,000	2,000 (1.50)	2,000 (1.00)	2,000 (1.00)	3,000 (3.33)	1,000	—
Thoropass	9,000	2,000 (5.00)	—	6,000 (2.00)	1,000 (4.00)	—	—
CATAAM	8,000	—	7,000 (1.43)	—	—	1,000	—
Scytale	8,000	—	2,000 (6.00)	6,000 (3.00)	—	—	—
Workiva	8,000	2,000 (8.00)	5,000 (4.40)	1,000 (4.00)	—	—	—
Live Compliance	8,000	—	7,000 (2.57)	1,000 (2.00)	—	—	—
Strac	7,000	—	—	1,000 (5.00)	—	—	6,000 (1.00)
Orbiq	7,000	—	—	—	—	—	7,000 (3.67)
Accountable HQ	7,000	—	4,000	2,000 (2.00)	—	1,000	—

Appendix B. Source mix and most-cited domains

Table B1. Source-type mix by AI engine — cited sources by classified source type; n = 3,231,000 cited sources

AI engine	Government	Wikipedia	Forum	Marketplace	Vendor site	Review aggregator	Social	Self	Other	Total
ChatGPT	56,000	7,000	12,000	2,000	407,000	18,000	6,000	—	1,172,000	1,680,000
Microsoft Copilot	—	—	—	—	—	2,000	—	—	150,000	152,000
Gemini	—	—	3,000	—	23,000	1,000	—	1,000	247,000	275,000
Google AI Mode	—	—	5,000	—	80,000	5,000	5,000	—	361,000	456,000
Google AI Overview	—	—	2,000	—	56,000	1,000	7,000	—	282,000	348,000
Perplexity	—	—	—	—	54,000	19,000	—	—	247,000	320,000
Total	**56,000**	**7,000**	**22,000**	**2,000**	**620,000**	**46,000**	**18,000**	**1,000**	**2,459,000**	**3,231,000**

Table B2. Five most-cited domains by AI engine — cited-source counts in parentheses; n = 3,231,000 cited sources

AI engine	Most-cited domains
ChatGPT	vanta.com (97,000), drata.com (89,000), archerirm.com (79,000), servicenow.com (77,000), gartner.com (76,000)
Microsoft Copilot	worldmetrics.org (38,000), riskpublishing.com (15,000), cataam.com (7,000), expertinsights.com (7,000), lucidityconsult.net (7,000)
Gemini	scytale.ai (21,000), riskwatch.com (21,000), enzuzo.com (18,000), sprinto.com (17,000), optro.ai (15,000)

AI engine	Most-cited domains
Google AI Mode	sprinto.com (25,000), vanta.com (24,000), complyjet.com (16,000), optro.ai (16,000), youtube.com (15,000)
Google AI Overview	vanta.com (22,000), drata.com (19,000), optro.ai (12,000), sprinto.com (11,000), youtube.com (10,000)
Perplexity	complyjet.com (18,000), vanta.com (18,000), scrut.io (18,000), strac.io (14,000), orbiqhq.com (13,000)

Appendix C. Product entries

Table C1. All product entries — 117 entries ordered by mentions; a dash where no mean rank was recorded; n = 240,000 responses

#	Product	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
1	Vanta	Vanta	111,000	2.8	0.71	6	4
2	Drata	Drata	97,000	3.0	0.70	6	4
3	MetricStream	MetricStream	62,000	2.3	0.80	6	4
4	Secureframe	Secureframe	59,000	3.4	0.65	6	4
5	Sprinto	Sprinto	57,000	3.2	0.71	6	4
6	Archer	Archer	47,000	3.1	0.76	6	4
7	IBM OpenPages	IBM	44,000	4.3	0.75	6	4
8	LogicGate Risk Cloud	LogicGate	43,000	4.4	0.71	6	4
9	Optro	Optro	35,000	4.0	0.74	5	4
10	ServiceNow IRM	ServiceNow	34,000	1.9	0.82	5	4
11	Compliancy Group	Compliancy Group	33,000	1.6	0.79	6	4
12	Riskonnect	Riskonnect	31,000	4.1	0.72	5	4
13	OneTrust	OneTrust	31,000	4.6	0.72	6	4
14	AuditBoard	AuditBoard	29,000	3.8	0.72	5	4
15	ServiceNow GRC	ServiceNow	22,000	3.1	0.75	4	4
16	MedTrainer	MedTrainer	18,000	3.4	0.77	4	4
17	Diligent One	Diligent	17,000	4.7	0.71	4	4
18	Hyperproof	Hyperproof	15,000	5.1	0.63	6	4
19	Scrut Automation	Scrut	13,000	3.5	0.71	3	4
20	MetricStream Connected GRC	MetricStream	12,000	2.6	0.80	3	4
21	SAP GRC	SAP	12,000	3.1	0.69	3	4
22	Archer	OpenText	11,000	2.8	0.76	4	4

#	Product	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
23	Abyde	Abyde	11,000	3.0	0.76	3	4
24	HIPAA One	HIPAA One	11,000	3.7	0.72	3	4
25	Paubox	Paubox	11,000	4.9	0.69	5	4
26	ServiceNow IRM/GRC	ServiceNow	10,000	1.2	0.89	3	4
27	ServiceNow Integrated Risk Management (IRM)	ServiceNow	10,000	1.9	0.81	3	3
28	Medcurity	Medcurity	10,000	2.0	0.85	5	3
29	IBM OpenPages with Watson	IBM	8,000	1.0	0.70	1	4
30	CATAAM	CATAAM	8,000	1.4	0.78	2	4
31	Strac Comply	Strac	8,000	1.6	0.78	2	4
32	Live Compliance	Live Compliance	8,000	2.5	0.89	2	4
33	Workiva	Workiva	8,000	5.3	0.64	3	4
34	MetricStream Risk	MetricStream	7,000	1.8	0.78	1	4
35	Optro (formerly AuditBoard)	Optro	7,000	3.7	0.80	3	2
36	Scytale	Scytale	7,000	3.6	0.74	2	3
37	Orbiq	Orbiq	7,000	3.7	0.60	1	4
38	Thoropass	Thoropass	6,000	4.0	0.67	3	3
39	Diligent IRM	Diligent	6,000	4.0	0.58	1	4
40	Diligent One Platform	Diligent	6,000	5.3	0.81	1	4
41	OneTrust GRC	OneTrust	6,000	6.0	0.65	3	4
42	Diligent Risk Management	Diligent	6,000	6.5	0.62	1	4
43	ServiceNow Integrated Risk Management	ServiceNow	5,000	1.0	0.90	2	4
44	MetricStream GRC	MetricStream	5,000	2.3	0.80	4	3
45	Proofpoint Compliance	Proofpoint	5,000	3.4	0.62	1	3
46	Accountable	Accountable	4,000	3.8	0.75	3	3
47	Scrut Automation	Scrut Automation	4,000	4.0	0.72	3	2
48	Clearwater Compliance	Clearwater	4,000	6.0	0.79	3	2
49	Accountable HQ	Accountable	4,000	5.7	0.53	1	4
50	Compliance One	Compliance One	3,000	1.0	0.90	1	3
51	ServiceNow GRC / IRM	ServiceNow	3,000	1.0	0.85	2	1
52	VelocityEHS Risk Management	VelocityEHS	3,000	—	0.50	1	2
53	Accountable HQ	Accountable HQ	3,000	2.0	0.70	2	2

#	Product	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
54	Veeva Vault QMS	Veeva	3,000	2.7	0.63	1	2
55	Scrut	Scrut	3,000	3.0	0.63	2	3
56	ComplyAssistant	ComplyAssistant	3,000	3.5	0.67	1	2
57	HealthStream	HealthStream	3,000	5.0	0.68	2	2
58	Onspring	Onspring	3,000	5.0	0.60	1	3
59	Optro (AuditBoard)	AuditBoard	3,000	5.7	0.62	2	3
60	LogicGate	LogicGate	3,000	6.0	0.58	2	3
61	Diligent	Diligent	3,000	6.0	0.55	2	2
62	Centraleyes	Centraleyes	2,000	1.0	0.90	1	1
63	Klarity Health	Klarity	2,000	1.0	0.80	1	1
64	The Guard	Compliancy Group	2,000	1.0	0.75	2	2
65	NAVEX One	NAVEX	2,000	1.5	0.80	1	2
66	Vendorica	Vendorica	2,000	2.0	0.70	1	2
67	Sentrix	Sentrix	2,000	2.0	0.45	1	1
68	Compyl	Compyl	2,000	3.0	0.75	2	2
69	Archer (RSA)	RSA	2,000	4.0	0.65	1	1
70	Scrutineer.ai	Scrutineer	2,000	4.5	0.55	1	2
71	Jotform Enterprise	Jotform	2,000	5.0	0.55	2	1
72	MedStack	MedStack	2,000	6.0	0.80	1	2
73	Resolver	Resolver	2,000	6.0	0.60	2	1
74	Mitratech Alyne	Mitratech	2,000	6.5	0.57	1	2
75	Salesforce Health Cloud	Salesforce	2,000	9.0	0.75	1	1
76	Spruce Health	Spruce Health	2,000	8.5	0.65	1	1
77	MetricStream GRC Platform	MetricStream	1,000	1.0	0.90	1	1
78	Files.com	Files.com	1,000	1.0	0.80	1	1
79	SecureSlate	SecureSlate	1,000	—	0.80	1	1
80	Oracle GRC	Oracle	1,000	—	0.80	1	1
81	Healthcare Compliance Pros	Healthcare Compliance Pros	1,000	—	0.70	1	1
82	Tugboat Logic	Tugboat Logic	1,000	—	0.70	1	1
83	RiskCloud	RiskCloud	1,000	—	0.60	1	1
84	RiskWatch	RiskWatch	1,000	—	0.50	1	1
85	Caspio	Caspio	1,000	—	0.50	1	1

#	Product	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
86	ServiceNow Integrated Risk Management (IRM / GRC)	ServiceNow	1,000	2.0	0.85	1	1
87	ComplyAura	ComplyAura	1,000	2.0	0.85	1	1
88	Raize Orion	Raize	1,000	2.0	0.80	1	1
89	Archer IRM	Archer	1,000	2.0	0.80	1	1
90	MetricStream ERM	MetricStream	1,000	2.0	0.70	1	1
91	Compliance Manager GRC	Compliance Manager GRC	1,000	2.0	0.70	1	1
92	Klarity Health	Klarity Health	1,000	2.0	0.70	1	1
93	HIPAAsite	HIPAAsite	1,000	2.0	0.70	1	1
94	Archer (RSA Archer)	RSA	1,000	3.0	0.80	1	1
95	MetricStream ConnectedGRC	MetricStream	1,000	3.0	0.80	1	1
96	ClearDATA	ClearDATA	1,000	3.0	0.70	1	1
97	Evidr	Evidr	1,000	3.0	0.60	1	1
98	RSA Archer (Archer IRM)	RSA	1,000	3.0	0.60	1	1
99	Ventiv IRM	Ventiv	1,000	3.0	0.60	1	1
100	CyberStrong	CyberSaint	1,000	4.0	0.80	1	1
101	Archer (formerly RSA Archer)	Archer	1,000	4.0	0.80	1	1
102	Archer GRC	Archer	1,000	4.0	0.80	1	1
103	Archer (formerly RSA)	Archer	1,000	4.0	0.80	1	1
104	A-LIGN	A-LIGN	1,000	4.0	0.70	1	1
105	Archer (RSA Archer Suite)	Archer	1,000	4.0	0.60	1	1
106	pTrackly	pTrackly	1,000	4.0	0.60	1	1
107	TrioCyber	TrioCyber	1,000	4.0	0.60	1	1
108	RSA Archer (by OpenText)	OpenText	1,000	5.0	0.80	1	1
109	TigerConnect	TigerConnect	1,000	5.0	0.70	1	1
110	NAVEX One Risk & Governance	NAVEX	1,000	5.0	0.65	1	1
111	IBM OpenPages (with watsonx)	IBM	1,000	6.0	0.75	1	1
112	Diligent One / ERM	Diligent	1,000	6.0	0.65	1	1
113	Total HIPAA	Total HIPAA	1,000	6.0	0.65	1	1
114	Virtru	Virtru	1,000	6.0	0.60	1	1
115	Healthicity	Healthicity	1,000	7.0	0.80	1	1
116	SimplePractice	SimplePractice	1,000	8.0	0.60	1	1

#	Product	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
117	Carepatron	Carepatron	1,000	9.0	0.50	1	1

Appendix D. Study parameters

Table D1. Study parameters — as configured for this study

Parameter	Value
AI engines	ChatGPT, Microsoft Copilot, Gemini, Google AI Mode, Google AI Overview, Perplexity
Markets	United States, Canada, India, Germany
Buyer-intent queries	10,000
Responses per engine	40,000
Responses in total	240,000
Corpus identity	10,000 queries × 6 engines × 4 markets = 240,000 responses
Query themes	GRC software selection; integrated risk management platforms; HIPAA compliance tooling; multi-framework compliance automation
Paraphrase variants	Four paraphrases of each core question
Frameworks referenced in the query set	SOC 2, ISO 27001, GDPR, PCI DSS, HIPAA
Collection window	August 2026
Cited sources recorded	3,231,000
Distinct product entries	117
Brands recorded	30

Appendix E. Metric computation

Table E1. Metric computation — the arithmetic behind each derived metric in this report

Metric	Computation
Responses per engine	queries × markets
Responses in total	queries × engines × markets
Mean response length	sum of response word counts ÷ responses
Hedging, recency-anchored, truncation and refusal rates	responses exhibiting the behaviour ÷ responses, as a percentage
Brand mentions, total	sum of that brand's per-engine mention counts

Metric	Computation
Mean ordinal position	sum of recorded positions ÷ mentions with a recorded position
Mean sentiment	sum of mention-level polarity scores ÷ mentions scored
Cited sources, total	sum of per-engine cited-source counts
Dead links, per engine	cited sources × dead-link rate, rounded
Dead links, total	sum of per-engine dead-link counts
Dead-link rate, overall	total dead links ÷ total cited sources, as a percentage
Distinct domains	count of unique registrable domains cited by that engine; never summed across engines
Source-type share	cited sources in the type ÷ that engine's cited sources; the residual “other” category carries the balance
Jaccard coefficient	$\|A \cap B\| \div \|A \cup B\|$ over the two engines' cited-domain sets
Mean pairwise Jaccard	sum of the 15 distinct off-diagonal coefficients ÷ 15
Product mentions	mentions of that product entry across the corpus
Mean rank	sum of recorded ranks ÷ mentions with a recorded rank